

POPULATION GENETICS PROJECT #6

Those of you who attended Nora's dissertation defense last Wednesday know that one chapter of her dissertation examined patterns of phenotypic variation in two closely related species of *Protea*, *P. punctata* and *P. venusta* occurring along a steep elevational gradient on Blesberg Mountain. What Nora didn't mention in her defense is that in addition to measuring plant traits in the field, she collected data on the number of seedheads produced by each individual — a good proxy of fitness in this group. You may also remember that she collected seed from many of the individuals measured in the field and grew progeny in the greenhouse.

You'll find two spreadsheets of data on the course website that include some of the data derived from these observations:

1. `field.csv` contains the field observations with one line of data for each individual. The first column is the species identity, the second is the number of seedheads, the third is leaf mass per area (`lma`), the fourth is leaf fresh water content (`fwc`), and the fifth is leaf length-width ratio (`lwr`).
2. `greenhouse.csv` contains the greenhouse observations with one line of data for each individual. The first column is the species identity, the second is a numeric index identifying the maternal plant for this individual.¹ The last three columns are `lma`, `fwc`, and `lwr`.

Using these data, answer the following questions:

1. What is the heritability of each trait in each species?
2. What is the selection differential for each trait in each species?
3. If the heritability estimate from the greenhouse study were applicable to the wild populations, what would the predicted response to selection be for each trait in each species?

¹Numbers are assigned sequentially within each species, so `mom_id` 1 in *P. punctata* refers to a different maternal plant than `mom_id` in *P. venusta*.

4. For what trait/species combinations, if any, is there some evidence that selection favors either higher or lower trait values? To be a little more specific, for what trait/species combinations is the posterior probability that the selection differential is positive (or negative) reasonably high, say greater than 75% (3:1 odds)?

Hints

- Be sure to include indications of uncertainty, i.e., credible intervals with your answers to the first three questions.
- Treat the maternal families in the greenhouse as if all of the offspring of a single mother are half sibs. Since both of these species are outcrossing, this should be a pretty good approximation.
- As with the full-sib analysis in the notes for Thursday's lecture, the within group component of the variance is equal to the total phenotypic variance minus the covariance within groups. Combining this hint with the last one will allow you to get expressions for V_a and V_e in term of the within- and among-maternal family variance components.
- Use the following function in R to get the posterior probability that a particular parameter is greater than (or less than) zero.

```
sign <- function(fit, par.to.report="S") {  
  tmp <- fit$BUGSoutput$sims.list[[par.to.report]]  
  lt_0 <- sum(tmp < 0)/length(tmp)  
  gt_0 <- sum(tmp > 0)/length(tmp)  
  cat(" P(", par.to.report, " < 0) = ", round(lt_0, 3), "\n", sep="")  
  cat(" P(", par.to.report, " > 0) = ", round(gt_0, 3), "\n", sep="")  
}
```

Call it as `sign(<name of a JAGS object>, <name of a JAGS parameter>)`. If you leave off the name of the parameter, it will look for a parameter named `S` by default.

- You'll want to normalize the individual fitness measurements by dividing all of them by the mean fitness in the population.
- You'll want to standardize the trait values (so you can compare intensities of selection on different traits if you're so inclined). Use this function for that purpose:

```

standardize <- function(x) {
  x_bar <- mean(x, na.rm=TRUE)
  s_dev <- sd(x, na.rm=TRUE)
  return((x - x_bar)/s_dev)
}

```

If `x_data` is a vector containing the trait data you want to standardize, use `standardize` as

```
x_standardized <- standardize(x_data)
```

Notice that when you report the response to selection, you'll be reporting it on the standardized scale. Don't worry about converting it back to the original scale.

- In the regression of fitness on a trait, suppose the fitness of individual i is $y[i]$ and the trait in that individual is $x[i]$, then you can write the regression like this:

```

y[i] ~ dnorm(mu[i], tau)
mu <- beta.0 + beta.1*x[i]

```

Use the following priors in the regression:

```

beta.0 ~ dnorm(0.0, 1.0)
beta.1 ~ dnorm(0.0, 1.0)
var.resid ~ dgamma(2.0, 2.0)
tau <- 1.0/var.resid

```

- The selection differential on a trait is given by

$$S = \beta_1 \left(\frac{V_p}{\bar{w}} \right) ,$$

where β_1 is the selection gradient, i.e., the slope of the regression, V_p is the phenotypic variance, and \bar{w} is the mean fitness. Since the last bit of advice was to divide all of the individual fitness measurements by the mean fitness in the population, $\bar{w} = 1$.